A Novel Methodology for Designing Linear Phase IIR Filters

R.Ramanathan[1], K.P.Soman[2]

[1]Lecturer, Department of ECE,
Email: r_ramanathan@ettimadai.amrita.edu
[2]Professor & Head, Center for Excellence in Computational Engineering and Networking (CEN)
Email: kp_soman@amrita.edu
Amrita Vishwa Vidyapeetham, Coimbatore, India

Abstract—**This paper presents a novel technique for designing an Infinite Impulse Response (IIR) Filter with Linear Phase Response. The design of IIR filter is always a challenging task due to the reason that a Linear Phase Response is not realizable in this kind. The conventional techniques involve large number of samples and higher order filter for better approximation resulting in complex hardware for implementing the same. In addition, an extensive computational resource for obtaining the inverse of huge matrices is required. However, we propose a technique, which uses the frequency domain sampling along with the linear programming concept to achieve a filter design, which gives a best approximation for the linear phase response. The proposed method can give the closest response with less number of samples (only 10) and is computationally simple. We have presented the filter design along with its formulation and solving methodology. Numerical results are used to substantiate the efficiency of the proposed method.**

Index Terms—Infinite impulse response, linear programming, filter design, linear phase filters

I. INTRODUCTION

The design of the infinite impulse response filter is often restricted to the specification of the magnitude response due to the fact that a linear phase response is not physically realizable. Some methods in the literature try to achieve a good approximation of the linear phase response, but results in extensive and complex hardware. Thus, the design phase involves huge computational resources as well as implementation phase requires extensive hardware. We intend to simplify the computation in design as well as minimize the hardware by just opting for a second order filter. We introduce a concept of using the frequency domain samples of the desired magnitude and phase response of the filter and using the same to obtain the filter coefficients to implement the filter. The digital filter can be modeled as a linear system represented by

$$\mathbf{Ax} = \mathbf{b} \tag{1}$$

Where, A is the matrix containing the sampled frequency responses (magnitude & phase) of the given filter [1]. The given magnitude and phase responses of the filter can be sampled at a convenient sampling frequency and the number of samples can be conveniently chosen. We will now formulate the filter design problem as a linear

programming problem and then we will redefine the same problem with some constraints for solving the linear programming model. The frequency domain responses of the filter (both magnitude and phase) are used to model the linear programming problem. Let us consider the standard expression [2] for the filter transfer function,

$$\frac{Y(z)}{U(z)} = \frac{b_0 + b_1 z^{-1} + \dots + b_N z^{-N}}{1 + a_1 z^{-1} + \dots a_D z^{-D}} \tag{2}$$

Where $U(z)$ is the z-transform of the input signal, $Y(z)$ is the z-transform of the output signal, and 'a' and 'b' factors are real valued coefficients. From the above transfer function, we can write the time-domain difference equation that implements the filter as,

$$y(k) = -a_1 y(k-1) - \dots a_D y(k-D)$$
$$+ b_0 u(k) + \dots + b_N u(k-N). \tag{3}$$

where the a and b coefficients are exactly the same as in the transfer function above, k is the time index, $u(k)$ and $y(k)$ are the current values of the input and output (respectively), $u(k-N)$ was the input value N samples in the past, and $y(k-D)$ was the output value D samples in the past. Generally, if the frequency response of a system at a frequency ω_1 is given in magnitude/phase form as $A_1 \angle \phi_1$ the output amplitude will be A_1 times the input amplitude and the output phase will be shifted an angle $\varphi1$ relative to the input phase when a steady-state sine or cosine wave of frequency $\omega1$ is applied to the system. For instance, if the input to the system described above at time instance 'k' is $u_1(k) = \text{Cos}(k\omega_1 ts)$, where t_s is the sampling period, then the output will be $y_1(k) = A_1 \text{Cos}(k\omega_1 ts + \varphi_1)$. The input and output values at any sample time can be determined in a similar manner. The final sampled values in the matrix form is,

$$
\begin{bmatrix} y_1(0) \\ y_2(0) \\ \cdot \\ \cdot \\ y_N(0) \end{bmatrix} = \begin{bmatrix} -y_1(-1) & \dots & -y_1(-D) & u_1(0) & \dots & u_1(-N) \\ -Y_2(-1) & \dots & -Y_2(-D) & u_2(0) & \dots & u_2(-N) \\ \cdot & & \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ -y_M(-1) & \dots & -y_M(-D) & u_M(0) & \dots & u_M(-N) \end{bmatrix} * \begin{bmatrix} a_1 \\ \cdot \\ a_D \\ b_0 \\ \cdot \\ b_N \end{bmatrix}
$$
$$\tag{4}$$

The above equation is similar to the system $\mathbf{A x} \approx \mathbf{b}$; as given by (1). In the above formulation, we have set the

current sample index as zero. Our aim now is to solve the above problem, and get the filter coefficients given by the vector '$\mathbf{x}$'. Standard methods use the inverse of the matrix 'A' which involves a heavy computation for huge matrices [3],[4],[5],[6],[7],[8]. If the inverse of the matrix does not exist, then Singular Value Decomposition can be used for Pseudo inverses [9]. To avoid this and make the procedure computationally simple, we introduce the concept of linear programming into this context. We now formulate the above as a simple optimization problem with few constraints.

II. PROBLEM FORMULATION

In the above model discussed, the matrix 'A' contains the sampled coefficients at different frequencies and the column vector '$\mathbf{x}$' contains the filter coefficients. The first 'D' values in '$\mathbf{x}$' represent the coefficients a_1, a_2, etc in denominator and the next 'N' values give the b_0, b_1, b_2, etc coefficients in numerator. The column vector '$\mathbf{y}$' is the desired output value (both amplitude and phase) at specific frequencies. Now we can move on to formulate the optimization problem to find the filter coefficients. We first aim to find the best possible solution by introducing a factor 'ε', which is vector containing allowable error. Then we try to reduce the error to get the best solution by minimizing this error parameter.

$$\mathbf{A}\mathbf{x} + \boldsymbol{\varepsilon} = \mathbf{y};$$
$$\Rightarrow \boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{A}\mathbf{x} \tag{5}$$

Now our goal is to frame a minimization problem to reduce the error. Here our variables are '$\mathbf{x}$' and '$\boldsymbol{\varepsilon}$'.

$$\underset{x,\mathcal{E},a}{\mathbf{Min}} \quad \sum_i \varepsilon_i + \sum_i a_i$$

S.t $\mathbf{A}\mathbf{x} + \boldsymbol{\varepsilon} = \mathbf{y};\ 0 \le \boldsymbol{\varepsilon} \le \mathbf{a.*y}$ and $\mathbf{a} \ge \mathbf{0};$ (6)

Instead of directly manipulating the error parameter, we can define two variables $\mathbf{w_1}$ and $\mathbf{w_2}$ (essentially vectors of length same as '$\boldsymbol{\varepsilon}$') such that

$$\boldsymbol{\varepsilon} = \mathbf{w_1} - \mathbf{w_2}; \tag{7}$$

So our equation (5) becomes,

$$\mathbf{A}\mathbf{x} + \mathbf{w_1} - \mathbf{w_2} = \mathbf{y}; \tag{8}$$

We also have a possibility to restrict the maximum value error to a limit say a fraction of the actual response value, so that relatively there is no adverse change in the overall response. Letting that positive factor to be '$\mathbf{a}$' that contains the fraction (eg: 0.01 or 1/100 [th]), where '$\mathbf{a}$' is also a vector.

$$0 \le \boldsymbol{\varepsilon} \le \mathbf{a.*y} \tag{9}$$

Here '$.*$' is a Matlab operation, which is an element by element multiplication. Rewriting the above constraint using $\mathbf{w_1}$ and $\mathbf{w_2}$,

$$0 \le \mathbf{w_1} \le \mathbf{a.*y}$$
$$0 \le \mathbf{w_2} \le \mathbf{a.*y} \tag{10}$$

Now the problem, which we need to solve, is given by

$$\underset{x,w_1,w_2,a}{\mathbf{Min}} \quad \frac{1}{C_1}\sum_i w_1 + \frac{1}{C_2}\sum_i w_2 + \sum_i a_i$$

$$\text{S.t} \quad \mathbf{A}\mathbf{x} + \mathbf{w_1} - \mathbf{w_2} = \mathbf{y};$$

$$0 \le \mathbf{w_1} \le \mathbf{a.*y} \ , \ 0 \le \mathbf{w_2} \le \mathbf{a.*y} \text{ and } \mathbf{a} \ge \mathbf{0}; \tag{11}$$

Here C_1 and C_2 are control parameters that can be adjusted to vary the depth of optimization of $\mathbf{w_1}$ and $\mathbf{w_2}$ respectively. By the above formulation, we try to reduce the value of error 'ε' along with the maximum permissible limit defined by '$\mathbf{a}$'. When we attempt to reduce the value of '$\mathbf{a}$', indirectly the maximum value that $\mathbf{w_1}$ and $\mathbf{w_2}$ can assume is reduced which in turn results in the value of 'ε'. The problem is a chain reaction one. That mutually reduces the above said parameters. This formulation of the optimization problem is easily solvable by any simple tool, even with a spreadsheet package like Microsoft Excel. This formulation results in reduced computational complexity and less number of frequency domain samples (eg: 8 or 16).

III. NUMERICAL RESULTS

We have taken few examples [10], [11], [12] of low pass; band pass and high pass filter and attempted to solve for the filter coefficients. The knowledge of desired frequency response (both magnitude as well as phase) is the only requirement. We have considered a second order filter in all the cases discussed with 10 sample values. The number of samples is left to the choice of the designer with only condition that minimum number of samples be 10 for a good approximation. The proper choices of sample points play an important role in the response approximation. The sampling is done at the rate of 1000Hz or 10^{-3} seconds. Since we have chosen a second order filter, we have N=D=2. We have solved the optimization problem using Microsoft Excel and the results are verified by using Matlab. The obtained coefficients are used to plot the magnitude and phase response and the results are verified.

A. Illustration 1:

A Low pass IIR Filter was designed for the given Specifications as shown in figure 1. The signal is sampled and the sample values corresponding to amplitude and phase are tabulated in Table 1. The filter problem is formulated as a linear problem using the above discussed formulation and is solved using Excel and the coefficients obtained are listed in Table 2, also the response obtained using Matlab is shown in figure 1.

219

TABLE 1: SAMPLED DATA :(10 POINTS) - LPF

Amplitude	Phase (radians)	Frequency (Hz)
1.0	-0.1	2.7400
1.0	-0.5	50.0400
1.0	-0.97	97.9600
1.0	-1.01	101.8700
0.9	-1.1	110.0000
0.6	-1.25	125.0000
0.5	-1.25	150.0000
0.3	-1.25	178.2800
0.1	-1.25	300.4200
0.001	-1.25	450.5600

TABLE 2: OBTAINED COEFFICIENTS - LPF

a_0	a_1	a_2	b_1	b_2	b_3
1.00	-0.8663	0.3141	0.2141	0.1122	0.1209

It can be inferred that the phase response exhibits a linear characteristics in our pass band which is 0 – 100 Hz. The phase is exactly '-1 radians' at 100 Hz and it exactly fits a straight line with slope angle 45 degrees. We are not concerned of the phase relation in the stop band, but the phase seems to fit a linear curve in stop band also. The small deviation in the response in the transition from the pass band to stop band can be rectified by increasing the number of sample points in this region. We have experimented with 10 samples and tried to prove that with just 10 samples, we are able to get a good approximation. Therefore, by suitably increasing the number form 10 to 16 or 32, the obtained response will exactly follow the desired one.
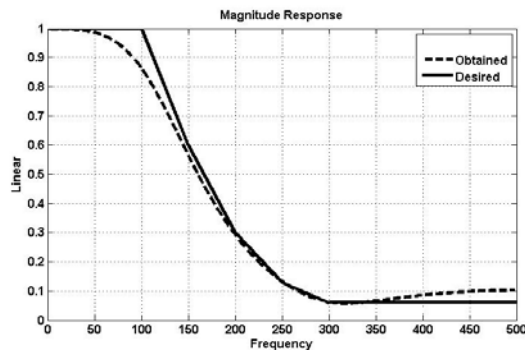

Fig1a. Comparison of desired and obtained Magnitude Responses of the low pass filter
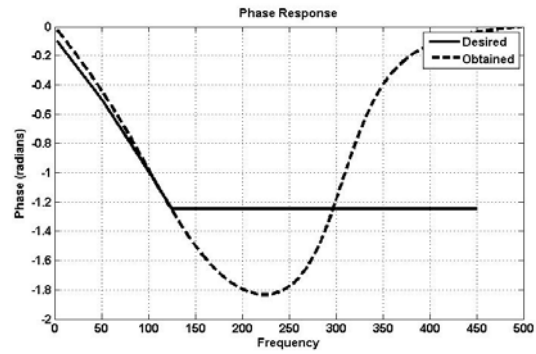

Fig 1b. Comparison of desired and obtained Phase Responses of the low pass filter

The desired and obtained magnitude and phase responses are comparatively studied and the figure 1 depicts the same. It is quite clear that the obtained response better approximates the desired one with just 10 samples and a second order IIR filter. With this method, a linear phase response can be achieved even with an IIR filter.

B. Illustration 2:

We now discuss the second example of designing a band pass filter to operate in the band from 100 Hz – 200 Hz with a linear phase response as depicted in fig 2b. The signal is sampled and the sample values corresponding to amplitude and phase are tabulated in Table 3. The filter problem is formulated as a linear problem using the above discussed formulation and is solved using Excel and the coefficients obtained are listed in Table 4 and the response obtained using Matlab is shown in figure 2.

TABLE 3: SAMPLED DATA (10 POINTS) - BPF

Amplitude	Phase (radians)	Frequency (Hz)
0.0	0.5	0.0000
0.25	0.5	50.0000
0.5	0.5	95.0000
0.9	0.5	100.0000
1.0	-0.11	150.0000
1.0	-0.63	200.0000
0.75	-1.0	250.0000
0.5	-1.46	300.0000
0.3	-1.5	350.0000
0.2	-1.5	450.5600

TABLE 4: OBTAINED COEFFICIENTS - BPF

220					
a_0	a_1	a_2	b_1	b_2	b_3
1.00	-0.7502	0.2644	0.2617	0.2471	-0.5010

The comparison plots (figure 2) between the desired and obtained responses are presented. The phase response in the pass band, that is 100 Hz – 200 Hz, is quite linear with respect to frequency. This region is of interest to us and is the key feature to be noticed.
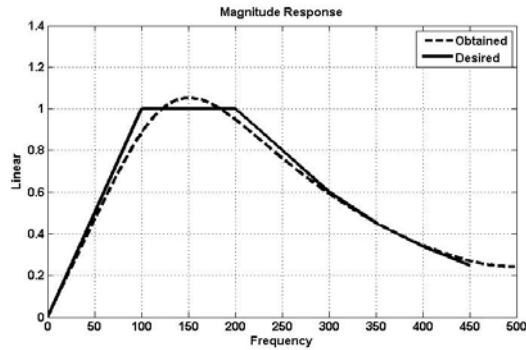

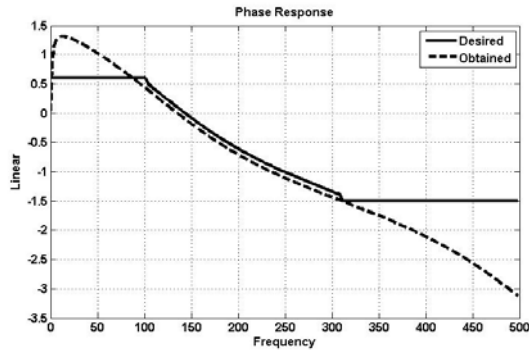Fig 2a. Comparison of desired and obtained Magnitude Responses of the Band pass filter

Fig 2b. Comparison of desired and obtained Phase Responses of the Band pass filter

C. Illustration 3:

The Third example we have considered is that of a high pass filter. The procedure similar to the previous examples is followed here also. The Desired responses (figure 3), sampled data (table 5), the obtained coefficients (table 6) and obtained responses (figure 3) are presented below. The comparison plot (figure 3) clearly describes the linear phase characteristics of the designed filter. By increasing the number of samples, we can make the obtained response exactly follow the designed one. Here we have focused on the linear phase response, hence chosen just 10 samples. By choosing 32 samples, the obtained response is found to be the exact copy of the desired response.

TABLE 5: SAMPLED DATA (10 POINTS) - HPF

Amplitude	Phase (radians)	Frequency (Hz)
0.0	1.7	0.0000
0.1	1.7	50.0000
0.2	1.7	95.0000
0.3	1.7	100.0000
0.5	1.5	150.0000
1.0	1.2	200.0000
1.0	0.9	250.0000
1.0	0.6	300.0000
1.0	0.4	350.0000
1.0	0.1	450.5600

TABLE 6: OBTAINED COEFFICIENTS - HPF

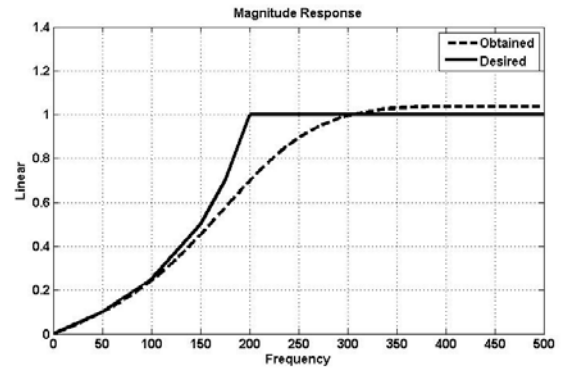| a_0	a_1	a_2	b_1	b_2	b_3
1.00	-0.2194	0.1662	0.4975	-0.7182	0.2208

Fig 3a. Comparison of desired and obtained Magnitude Responses of the high pass filter
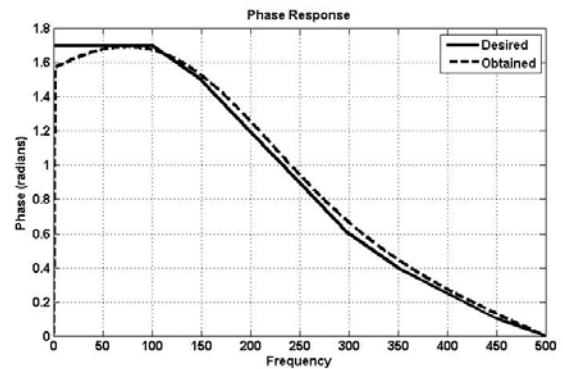

Fig 3b. Comparison of desired and obtained Phase Responses of the high pass filter

IV. PARAMETERS INFLUENCING THE FILTER DESIGN

Firstly, the filters designed by the above method solely depend upon the number of samples and the sample

values. The key point is the choice of the sample values. The frequency points where there is a significant change in the response either magnitude or the phase should be captured by the sample values we choose. For a normal filter, we have considered (low pass or high pass or band pass); four points in the transition band, three points in the pass and stop band will do the trick. However, if the filter should provide specific responses, then the number of samples should be chosen accordingly depending upon the level of approximation we require. However, as a general case a filter designed by this proposed methodology with 10 samples should yield a good approximation of the magnitude and phase response when compared to the design by existing methods. In addition, we can achieve a linear phase Infinite Impulse Response (IIR) filter which is normally not realizable. Hence, this method can be effectively used to achieve the desired response using an IIR filter.

Secondly, we are interested only in the phase response in the pass band and the phase in the stop band is not of our concern. Yet, the literatures reveal that the phase change in stop band will influence the overall response of the filter response. These gives a cushion of changing the phase in the stop band and further achieve better results. Therefore, we can introduce few more variables that are phase values in stop band in our formulation, even provide some lower and upper bounds, and further refine the formulation.

V. CONCLUSION AND FUTURE SCOPE

The formulation of the IIR filter design problem, with both the magnitude and phase specifications, is easily done using this proposed technique of linear programming. This application of linear programming for design of linear phase IIR filters is a novel methodology. This requires only very low computational resources in spite of a very good approximation of the linear phase response. The results obtained by the proposed method are quite encouraging. The results can be further improved by proper choice of the number of samples and position of sample points. These results presented above guarantee the efficiency of the proposed method. The literature shows that by slightly tuning the phase in the stop band (which is of no significance) the phase response in the pass band can be changed accordingly.

These phenomena can be incorporated in the proposed technique to achieve best results. The combination is implemented and the performance is being investigated. It is very much conducive that the proposed method may be effectively adopted by the design engineers. The proposed method proves to be a simple method when compared to the counterparts for providing better results.

REFERENCES

[1] G.Berchin,"Precise Filter Design" IEEE Signal Processing Magazine, January 2007.
[2] R. Lyons, *Understanding Digital Signal Processing*, 2nd ed., Upper Saddle River, NJ:Prentice-Hall, pp. 232–240, 2004.
[3] C. Sidney Burrus,"Direct Frequency Domain IIR Filter Design Methods"Version 1.1: Jun 24, 2008
[4]Ricardo A. Vargas and C. Sidney Burrus. On the design of IIR filters with arbitrary frequency response. In Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing , ICASSP-01, Salt Lake City, May 7-11 2001.
[5] C. Sidney Burrus and Ricardo A. Vargas. The direct design of recursive or IIR digital filters. In Proceedings of International Symposium on Communications, Control, and Signal Processing, ISCCSP-08, Malta, March 2008.
[6] L. B. Jackson. Frequency-domain steiglitz-mcbride method for least-squares IIR filters design, arma modeling, and periodogram smoothing. IEEE Signal Processing Letters, 15:498211;52, 2008.
[7] Takao Kobayashi and Satoshi Imai. Design of IIR digital filters with arbitrary log magnitude function by WLS techniques. IEEE Trans. on ASSP, 38(2):2478211;252, February 1990.
[8] Atmadji W. Soewito. Least Square Digital Filter Design in the Frequency Domain. PhD. thesis, Rice University, ECE Department, Houston, TX 77251, December 1990.
[9] G.Strang, "Linear Algebra and its Applications",3rd ed.,Thomson Brooks/Cole.
[10] S.K.Mitra, J. Kaiser, *Handbook for Digital Signal Processing*, 1993 John Wiley and Sons, Inc.
[11] T. W. Parks and C. S. Burrus. Digital Filter Design. John Wiley & Sons, New York, 1987
[12] L. R. Rabiner and B. Gold. Theory and Application of Digital Signal Processing. Prentice-Hall, Engle-wood Cliffs, NJ, 1975.